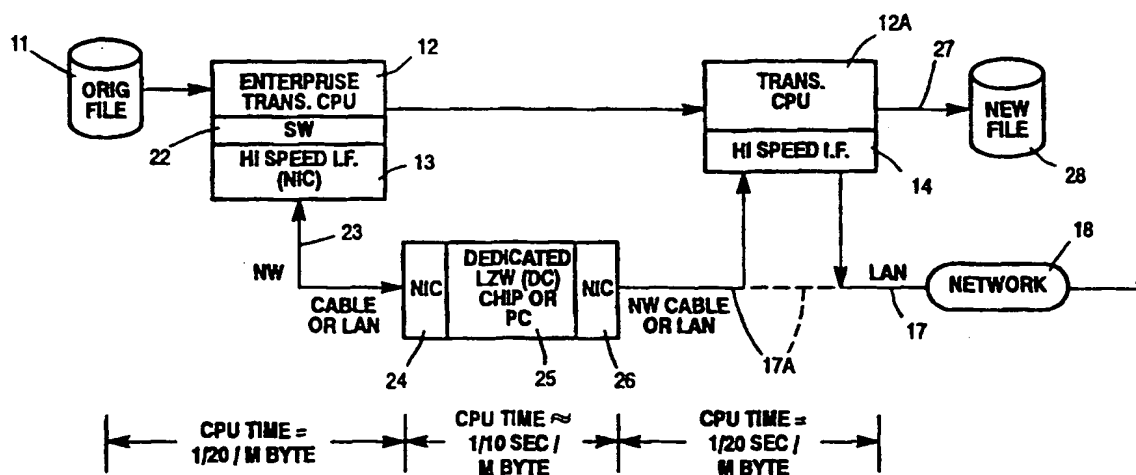




## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>G06F 9/38</b>		<b>A1</b>	(11) International Publication Number: <b>WO 99/67706</b>
			(43) International Publication Date: 29 December 1999 (29.12.99)
(21) International Application Number: PCT/US99/12405 (22) International Filing Date: 3 June 1999 (03.06.99) (30) Priority Data: 09/104,115          24 June 1998 (24.06.98)          US (71) Applicant: UNISYS CORPORATION [US/US]; Township Line and Union Meeting Roads, P.O. Box 500, Blue Bell, PA 19424-0001 (US). (72) Inventor: YORK, Kenneth, Lindsay; 2130 Country Club Drive, Huntington Valley, PA 19006 (US). (74) Agents: STARR, Mark, T. et al.; Unisys Corporation, Township Line and Union Meeting Roads, P.O. Box 500, Blue Bell, PA 19424-0001 (US).		(81) Designated States: JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  Published With international search report.	

(54) Title: SYSTEM FOR HIGH SPEED CONTINUOUS FILE TRANSFER PROCESSING



## (57) Abstract

In a system for economically and continuously transmitting data files (@11) from one location to a remote location (@29) there is provided a high speed mainframe computer (12) adapted to read a data file (11) and to determine if the process to be performed on the data requires a large amount of mainframe processor cycle time. If so the data file (11) is downloaded to a memory (@12) accessible by a much less expensive single chip type dedicated computer (25) that is particularly well suited to perform the process that would require high mainframe processor time and cost. The output of the dedicated computer (25) is coupled to a communication network (17) for transmittal to the remote location (28). In the preferred embodiment of the present invention first-in-first-out (FIFO) data queues (33) are employed to maintain optimum utilization of the dedicated computer (25) and transfer of the data files (11).

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

A SYSTEM FOR HIGH SPEED CONTINUOUS FILE TRANSFER PROCESSING  
5 OF DATA FILES

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to the transfer of data  
10 files from a first location to a second remotely located  
data file. More particularly, the present invention  
relates to a high-speed cost-effective data file transfer  
system.

15 2. Description of the Prior Art

Systems for transferring data to a centrally located  
enterprise system from remotely located areas over various  
links or networks are well known in the prior art and are  
classified in class 370, subclasses 260+, 465+ and 536+.  
20 One such system is shown and described in my U.S. patent  
5,680,400 which is assigned to the same assignee as the  
present invention and is incorporated by reference herein.  
One purpose of this system is to split the data stream into  
a plurality of parallel streams for an asynchronous data  
25 transfer over slow parallel networks or links for the  
purpose of speeding up the total flow of data over slow  
networks. When plural slow networks are accessible, the  
division of the data into plural parallel streams can speed  
up the flow of data transfer. Over long routes, plural  
30 slow networks may appear in the overall transmission paths.

being used and other means for speeding up the data transfer may be employed.

There are other bottle necks to be found in data transmission systems which were heretofore ignored. Many large general purpose mainframes computer systems often employ expensive transaction processing computers. Such expensive transaction processing computers serve their main purpose very well but when used to perform certain other tasks do not perform as well as other much less expensive one chip computers. When such a mainframe computer is in an enterprise system, a multiple processing system or a stand alone configuration and is also used to perform these certain other "difficult" tasks, they are performed at a higher economical cost than is necessary. Further, when the mainframe computer is functionally placed in a series path of the data transmission flow, the rate of data transfer is attenuated by the increase computational time cost of the slower computer.

An example will illustrate the point. The time required for a typical large expensive mainframe transaction computer to perform a certain process and formatting tasks have been measured to be between 1.5 and 2 seconds of computer processing per mega byte of data where the identical process has been and is capable of being performed in less than  $1/20^{\text{th}}$  of this time by use of a dedicated one chip general purpose processor.

Accordingly it will be highly desirable to provide in an expensive mainframe computing system a method and apparatus for identifying these certain difficult and highly processor cycle intensive tasks and providing means

for off loading such difficult tasks to a dedicated microprocessor or single chip computer in a manner which unburdens the mainframe computer of the difficult task without a time cost penalty.

5

#### SUMMARY OF THE INVENTION

It is a primary object of the present invention to provide a computing system having at least one high cost mainframe computer coupled to a low cost high speed  
10 dedicated computer chip or microprocessor or PC for performing difficult processing tasks which run much faster on the dedicated computer.

It is another primary object of the present invention to provide a low cost accelerating processor coupled to a  
15 high cost mainframe computer for performing certain difficult tasks that are capable of being performed on the mainframe computer at high processor time costs.

It is another primary object of the present invention to provide in an expensive mainframe computer system means  
20 for determining difficult time consuming tasks which can be run both faster and cheaper on a dedicated and low cost computer.

It is another primary object of the present invention to provide in a computing system a method and means for  
25 off-loading difficult tasks assigned to one computer to a dedicated computer which can more efficiently run a difficult task.

It is the general object of the present invention to provide a processing system for continuously processing a  
30 stream of data extracted from a data file through two

different types of computers to while significantly enhancing the speed or rate of processing the stream of data.

According to these and other objects of the present invention there is provided a mainframe computer for  
5 accessing data from a data file and for producing a data stream to be processed and transmitted to a user source in a most efficient manner. A low cost dedicated processor is coupled in the output data stream of the mainframe  
10 computer. The process to be performed by the mainframe computer is examined and that stage of the process which is identified as a difficult process is not performed on the data stream in the mainframe computer but it performed downstream by the dedicated processor, whereby the total  
15 process to be performed by the mainframe computer is performed both faster and cheaper by the combination of the two computers. The cost of the downstream computer is insignificant in comparison compared to the mainframe computer.

20

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic block diagram of the prior art mainframe computer showing the CPU time required to process 1.0 mega bytes in a data stream;

25 Figure 2 is a schematic block diagram of the prior art mainframe computer of Figure 1 showing the CPU processing time required to impose data compression process on the data streams;

Figure 3 is a schematic block diagram of the present  
30 invention system showing a continuous flow of data in a

data retrieval and transport system having a low cost dedicated processor connectable in one of the data pipeline paths when needed to perform difficult tasks; and

Figure 4 is a schematic block diagram showing the steps employed in the data pipeline or data stream to achieve a continuous flow of data in an asynchronous mode of operation.

#### DETAIL DESCRIPTION OF THE PREFERRED EMBODIMENT

10 Refer now to Fig. 1 showing a block diagram of a prior art mainframe computer system 10 which comprises a memory file 11 that is coupled to a data retrieval CPU 12. CPU 12 is shown provided with an interface 13 which is coupled to an interface 14 on a data transmission CPU 15 which may be the same as CPU 12. CPU 15 is shown provided with an interface 16 which couples to a local area network 17 or other link which connects to a network 18. It will be understood that the network 18 may be coupled to a remote computing system located almost anywhere in the world.

20 A feature of the data retrieval computer 12 is that it is capable of retrieving data and transporting data at an extremely high rate of speed. Fig. 1 shows that the CPU time for retrieving one mega byte of data is approximately  $1/20^{\text{th}}$  of 1 second. When this data is retrieved it is in the form of a data stream and may be transmitted directly through an interface 13 to the data transaction CPU 15 which may be the same computer 12 and its transaction time for each mega byte of data in the data stream is also approximately  $1/20^{\text{th}}$  of 1 second. Thus the overall time to  
25  
30 extract or retrieve the data from the memory file 11 and

present it to the LAN network 17,18 is approximately  $1/10^{\text{th}}$  of 1 second.

Refer now to Fig. 2 showing a schematic block diagram of a prior art mainframe computer of the type shown in Fig. 1 where the mainframe computer is used to perform a data  
5 compression process on the data stream. The modified computing system 10A is shown having the identical memory file 11 coupled to a data retrieval mainframe computer 12. The computer 12 is coupled by a software connection 19 to  
10 the same data retrieval computer 12 which is programmed to perform a Lempel/Ziv/Welch (LZW) data compression algorithm. The computer 12 which performs the data compression process is software coupled via link 21 to the same data transmission computer 12 as explained above.  
15 Computer 12 is provided with an interface 13 which couples to a LAN network 17 and wide area network 18.

The same computing time of  $1/20^{\text{th}}$  of 1 second is shown for the computer 12 to retrieve data from the file 11. The same transaction time of  $1/20^{\text{th}}$  of 1 second is shown for the  
20 data transport process of CPU 12. However, interposed between data retrieval and data transport processes there is shown a time of 1.5 - 2 second per mega byte of data in the data stream to perform a data compression operation because of the computer cycle intensive algorithm (e.g.  
25 LZW). Thus the overall computing time of approximately 1.6 to 2.1 seconds per mega of data passing in the data stream is consumed when the mainframe computer is employed to perform all three processes noted.

Refer now to Fig 3 showing a schematic block diagram  
30 of the preferred embodiment of the present invention. The



system 20 is shown having a continuous flow path for data from a original file 11 into an enterprise transaction computer 12 which is provided with a high speed interface 13. Further, the transaction computer is provided with a software program 22 capable of determining when a difficult process should be off loaded to a dedicated computer. The data stream created by the CPU 12 is shown being transferred across a connection 23 which may be a cable, an internal bus or an external network connection that is coupled into an interface 24 of the dedicated computer 25. Computer 25 performs the difficult process and has a second or the same interface 26 which connects to a connection 17A which again may be a cable or a network interface connection. The internal bus 17A is shown connected to the high speed interface 14 of the transaction computer 12A which performs the transport process and outputs the continuous data stream to a connection 27. For purposes of the present illustration the connection 27 is a cable or link connection to a local new file 28. It will be understood that the same data stream may be outputted to the network 17, 18 which delivers the data stream to a file server (not shown) for producing a new file anywhere on the network. To compare Fig. 3 to the system shown in Fig. 2 and illustrate the present improvement a connection 17A is shown as a phantom line which connects via CPU 12A to the network 18 as explained hereinbefore. Thus, it will be understood that the dedicated computer 25 can deliver the stream of processed data directly to the network 18 for storage at remote new files located anywhere on the network.

Refer now to the CPU time shown as  $1/20^{\text{th}}$  of 1 second per mega byte for delivering the data stream to the interface 24. Also note that the dedicated computer time for processing the difficult process is very fast and may be as small as  $1/10^{\text{th}}$  of 1 second per mega byte for performing the LZW data compression algorithm. Further, CPU transaction time is shown as  $1/20^{\text{th}}$  of 1 second per mega byte and represents more than the transaction time required for the transaction process at 12A and would require even less time if performed by the dedicated processor 25. The total time for performing the same process performed in less than 1.6 seconds in Fig. 2 is now shown to be performed in at least approximately in  $1/5^{\text{th}}$  of 1 second thus unburdening the transaction processor 12 or 12A for other work. In the preferred embodiment of the present invention the dedicated processor may be a low cost Intel processor having a 200 mega hertz internal clock with Pentium II <sup>™</sup> capabilities. Other processors for similar or higher performance costs well below \$2000 and will perform the data compression process at higher speeds. When it is considered that the mainframe computers may cost several millions of dollars, the added cost of the dedicated processor is an insignificant item and more than pays for itself in conserving valuable mainframe computer time.

Refer now to Fig. 4 showing a schematic block diagram of steps that are employed in the data pipeline in Fig. 3 to achieve a continuous flow of data in an asynchronous mode of operation. An original file 11 is shown as a disk but may be any other form of memory which is accessible at high speeds from which data can be retrieved by the CPU 12

in a step labeled retrieval 29. After retrieving the information from the original file 11 it is placed in a first in first out (fifo) data queue shown as block 31. A LZW data compression algorithm is performed at step 32 in a  
5 continuos nonstop mode using the information supplied by the queue 31. The data processed at the processing step 32 is then passed to a send fifo queue 33 to buffer the information before it is supplied to the data transaction CPU 12A to perform the data transmit step 34. Then CPU 12A  
10 can place the information on the LAN 17 for transmission over the network 18. As explained hereinbefore the data stream can be directly connected to a local file 28 for later use.

It will be appreciated that in some data transfer and  
15 file transfer systems the user or customer desires to create a data compressed archival file 28 for later use. Some customers desire to conserve memory space by performing the data compression techniques which are described in our U.S. patents 4,558,302 and 4,971,407.  
20 Such data compression techniques may compress the data to a format of one-half to one-seventh of the space previous used in a format which is data loss less.

Having explained a preferred embodiment of the present invention and alternatives or modification for the use of  
25 the system, it will be appreciated that the mainframe computer 12 usually has advanced knowledge of the type of processes it will be requested to perform. Thus, when the process is noted that becomes difficult (or has subprocess steps that are difficult) the mainframe computer 12  
30 recognizes this, and is programmed at block 22 to down load

or off load the difficult process onto a dedicated high speed processor 25 of the type described hereinbefore. Further, some difficult processes that may be performed more efficiently by the dedicated processor 25 will not  
5 process the data stream supplied at connection 23 as fast as the mainframe computer 12 can retrieve it from its original file. In this case the retrieval step 29 shown in Fig 4 is operated in an intermitted fashion so as to maintain the fifo queue 31 full without the dedicated  
10 processor 25 knowing the speed at which the information is being retrieved. Thus, the mainframe computer 12 is free to perform other operations while intermittently loading the queue 31.

Having explained a preferred embodiment method and use  
15 of the present invention it will be appreciated that other difficult processes than data compression can be performed in the same manner to achieve economies of processing in a the mainframe computer system virtually without cost and the invention described herein is only limited by the  
20 claims appended hereto.

CLAIMS

1. A method for decreasing the overall time for transmitting a data file (11) onto a network, comprising  
5 the steps of:

reading data from a data file (11) at high speed employing a mainframe computer (12),

10 determining (@12) that a difficult process to be performed on the data requires more than twice the time required to read said data from said data file,

down loading said data to a memory (@12 or  
15 25) accessible by a single chip type dedicated computer (25),

performing said difficult process in said dedicated computer (25) programmed to perform the process,  
20

generating a stream of processed data (@17) representative of said data file at an output port of said dedicated computer (25), and

25 coupling said stream of processed data to a communication network for transmittal to a new file (28).

2. A method as set forth in claim 1 wherein said step of determining (@12) is predetermined for said  
30 difficult process to be performed.

3. A method as set forth in claim 2 wherein said difficult process comprises one or more data compression processes.

5

4. A method as set forth in claim 1 wherein said step of reading data is performed at speeds in excess of said dedicated computer to perform said difficult process, and

10

said step of down loading data to a memory (12 or 25) comprises intermittently down loading data to the accessible memory to maintain a data supply to an input of said dedicated computer (25).

15

5. A method as set forth in claim 4 wherein said accessible memory comprises a first-in-first-out (FIFO) data queue.

20

6. A method as set forth in claim 5 wherein said FIFO data queue is mounted in said mainframe computer.

25

7. A method as set forth in claim 1 wherein said dedicated computer comprises a motherboard and associated memory mounted in said mainframe computer.

30

8. A method as set forth in claim 7 wherein said dedicated computer comprises a low cost personal computer (PC) programmed to perform data compression and to produce said stream of processed data.

9. A method as set forth in claim 8 wherein said output port of said dedicated computer is coupled to a second first-in-first-out (FIFO) data queue.

5

10. A method as set forth in claim 9 wherein said second FIFO is coupled to said mainframe computer and further includes,

10

writing said stream of processed data to a new file.

15

11. A method as set forth in claim 9 wherein said second FIFO is coupled to said mainframe computer and further includes,

writing said stream of processed data on to a network communication line for transfer to a remotely located data file.

20

12. A method as set forth in claim 10 which further includes,

reading said new file, and

25

writing said new file onto a network communications line for transfer to a remotely located data file.

13. A method as set forth in claim 10 wherein said mainframe computer comprises means for reading and writing of data transactions at high speeds.



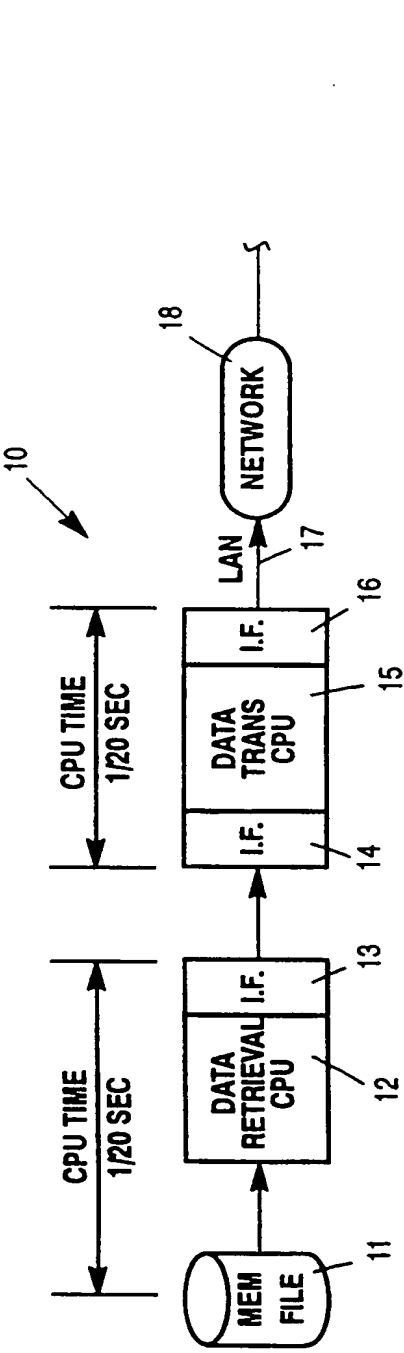


Figure 1 (Prior Art)

10A

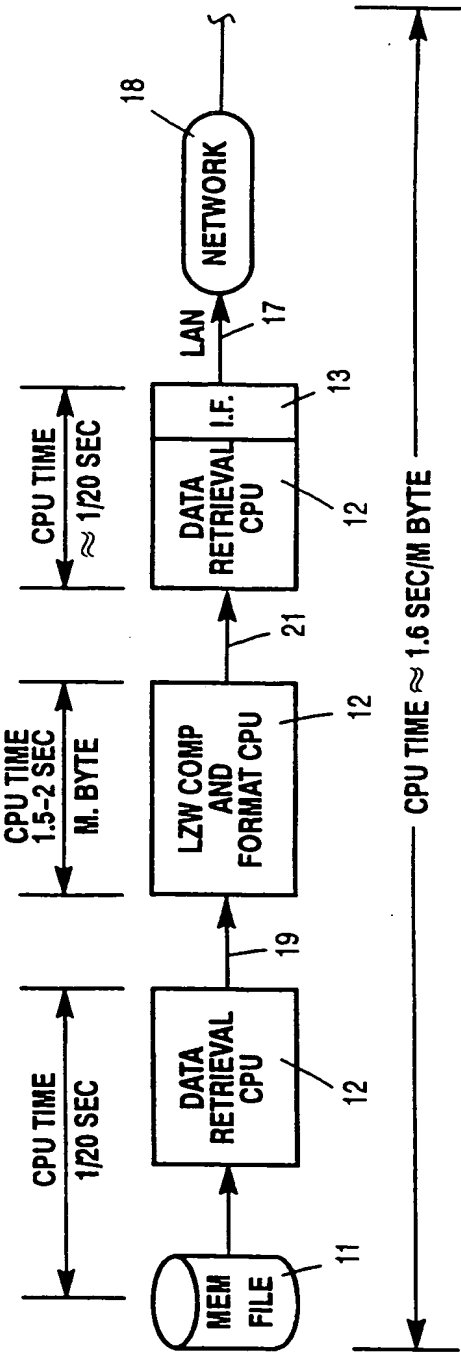
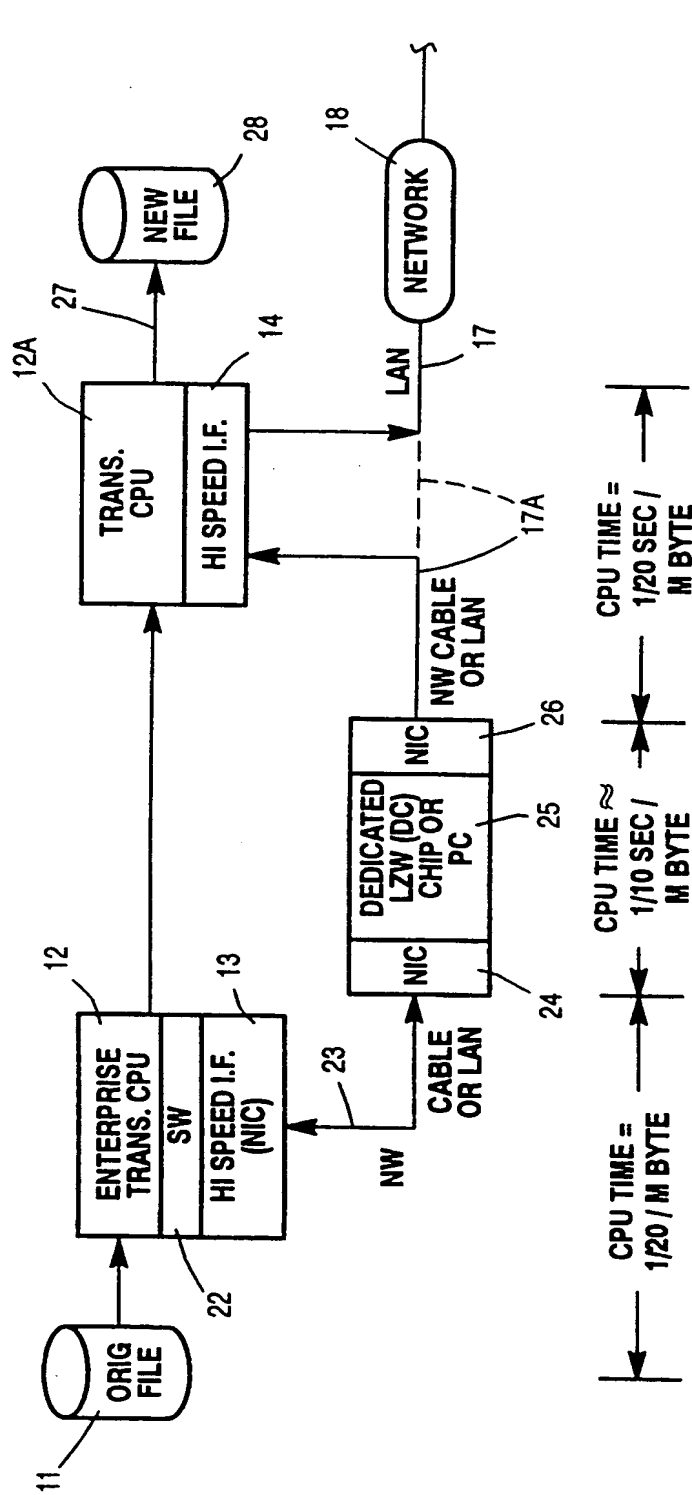
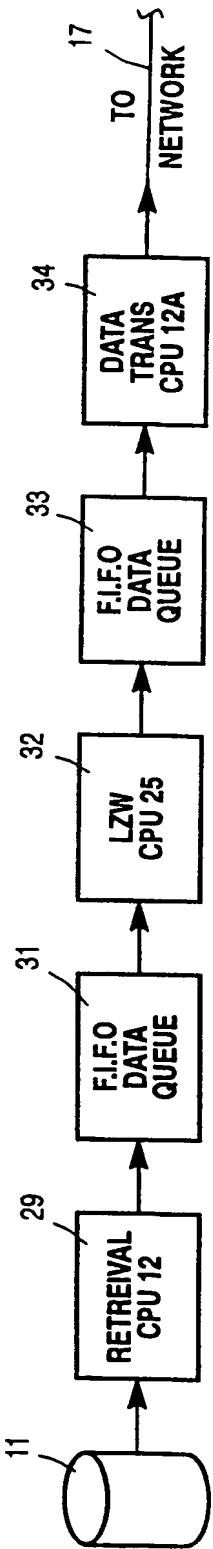


Figure 2 (Prior Art)



**Figure 3**



**Figure 4**

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 G06F9/38

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 357 614 A (PATTISAM ET AL) 18 October 1994 (1994-10-18) column 1, line 21 -column 2, line 10 column 2, line 25 -column 3, line 36 column 10, line 61 -column 12, line 36 ----	1-13
A	US 5 706 489 A (CHI ET AL) 6 January 1998 (1998-01-06) column 1, line 50 -column 2, line 32 column 2, line 48 - line 55 column 5, line 13 -column 30 ----	1-13
A	US 5 655 146 A (BAUM ET AL) 5 August 1997 (1997-08-05) column 1, line 63 -column 2, line 7 -----	1-13



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

## \* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&amp;" document member of the same patent family

Date of the actual completion of the international search

4 October 1999

Date of mailing of the international search report

11/10/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Abram, R

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/12405

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5357614 A	18-10-1994	NONE	
US 5706489 A	06-01-1998	NONE	
US 5655146 A	05-08-1997	CA 2137488 A,C EP 0668560 A JP 2633488 B JP 7239783 A	19-08-1995 23-08-1995 23-07-1997 12-09-1995